The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Experience-based Learning Mechanism for Neural Controller Adaptation: Application to Walking Biped Robots

John Nassour, Patrick Hénaff, Fethi Ben Ouezdou and Gordon Cheng

Abstract— **Neurobiology studies showed that the role of the Anterior Cingulate Cortex of the brain is primarily responsible for avoiding repeated mistakes. According to vigilance threshold, which denotes the tolerance to risks, we can differentiate between a learning mechanism that takes risks, and one that averts risks. The tolerance to risk plays an important role in such learning mechanism. Results have shown the differences in learning capacity between risk-taking and risk avert behaviors. In this paper, we propose a learning mechanism that is able to learn from negative and positive feedback. It is composed of two phases, evaluation and decision-making phase. In the evaluation phase, we use a Kohonen Self Organizing Map technique to represent success and failure. Decision-making is based on an early warning mechanism that enables to avoid repeating past mistakes. Our approach is presented with an implementation on a simulated planar biped robot, controlled by a reflexive low-level neural controller. The learning system adapts the dynamics and range of a hip sensor neuron of the controller in order for the robot to walk on flat or sloped terrain. Results show that success and failure maps can learn better with a threshold that is more tolerant to risk. This gives rise to robustness to the controller even in the presence of slope variations.**

I. Introduction

Some cognitive studies have identified an early warning system in the human brain that can avoid to make past mistakes again. They have shown how the brain remembers details about past dangers [1]. An activity was found in the Anterior Cingulate Cortex (ACC) after making mistakes [2]. This cortex area works as an early warning system that adjusts the behavior to avoid dangerous situations. It responds not only to the sources of errors (external error feedback), but also to the earliest sources of error information available (internal error detection) [3]. It becomes active in proportion to the occurrence likelihood of an error [4][5][6]. Therefore, it can learn to identify situations where humans may make mistakes, and then help to avoid such situations to occur [2]. It learns to predict error likelihood even for situations where no error occurs previously [7]. Through the observation of particular areas located in cerebral cortex in the brain responsible for cognitive control, neuropsychological studies demonstrated a switching in human learning strategies around the age of twelve years. This switch from learning with positive feedback to learning with negative feedback probably comes from the combination of brain maturing and experience [8].

F.B. Ouezdou and J. Nassour are with Versailles Saint Quentin University-France ({fethi.benouezdou, nassour}@lisv.uvsq.fr).

P. Hénaff is with Versailles Saint Quentin University and Cergy Pontoise University-France (patrick.henaff@uvsq.fr).

G. Cheng is with Technical University-Munich (gordon.cheng@ieee.org).

Our work aims to produce an early warning mechanism that can help to avoid repeating past errors in the generation of walking patterns for humanoid robots. It is necessary for such mechanism to have an experience in mistakes and other experience in success, in order to evaluate new situations before taking any decision and carrying out the test on the robot. This mechanism of selection allows to determine the state space of parameters in the zone of success and also in the zone of conflict. It is used to adapt the dynamics and range of a hip sensory neuron in a neural reflexive controller, proposed by Wörgötter [9], for simulated planar biped robot in order to avoid falls when the slope of terrain varies.

This paper is structured as follows. The second section presents the principles of our learning mechanism, and introduces the concept of vigilance. The third section describes the neural reflexive controller based on sensory motor neurons [9]. In the fourth section we use this mechanism to detect the domain of viability of the controller for walking on flat terrain and then on sloped terrain, where the effect of vigilance threshold on learning was investigated. Finally, we conclude this paper with some research perspectives.

II. Learning Mechanism

The objectives of this learning mechanism is to adapt parameters of a low level controller and detecting its domain of viability, which brings more adaptation to external and internal perturbations. We designate by V the state space of those parameters. The mechanism must be able to learn from negative feedback (failure) and positive feedback (success). Therefore, it must have experience with success and other with failure in the state space V. As each vector $\overrightarrow{v}$ from V leads to either success or failure, the mechanism will evaluate whether this vector belongs to the success domain or to the failure domain. The decision mechanism ("go", "nogo"), described in [10], works as an early warning system similar to that in the Anterior Cingulate Cortex [2][7]. The learning architecture is then based on these two mechanisms and works as shown in Fig. 1.

A. Success-failure evaluation

To represent the knowledge in success and in failure, we define two independent neural networks that are well-known Self Organizing Maps, proposed by Kohonen [11]. Success map learns in case of success trials, and failure map learns in case of failure trials. During the learning, the two maps will be self-organized in the state space that will be therefore divided into three zones: a zone of success represented by success map, a zone of failure represented by failure map,

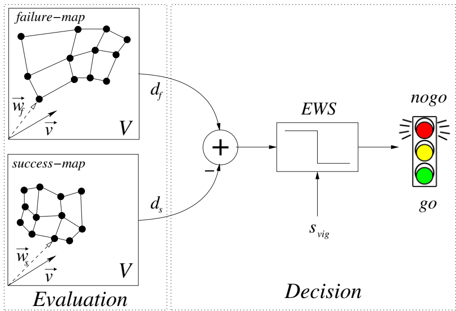


Fig. 1. Learning mechanism with evaluation and decision phases.

and a zone of conflict that corresponds to the interference between the two maps. The evaluation of any vector $\overrightarrow{v}$ from space V belonging to success or failure is defined by the distance between $\overrightarrow{v}$ and each map. The distance of a vector with a map is the distance between this vector and the neuron that is closest to it in the state space (the neuron winner). For each $\overrightarrow{v}$ we have therefore two distances: one to success map called d_s, and another to failure map called d_f.

B. Decision mechanism

For a vector $\overrightarrow{v}$, the comparison between the distance with success map d_s and the distance with failure map d_f leads to an expected result in the case where the vector is applied on the low level controller (trial). According to expected result, if it may lead to failure, then an Early Warning Signal EWS becomes active to avoid the passing into the lower level, and the decision will be "nogo". When EWS is inactive the decision is "go". The decision mechanism is affected by the threshold of vigilance s_{vig}, which will be detailed later.

C. Learning algorithm

Success and failure maps represent the knowledge in success and in failure inside the state space. First maps will be initialized in the state space V. Then, we take one vector $\overrightarrow{v}$ randomly from this space. In the phase of evaluation, we calculate the distance between this vector and all the neurons of both maps, as in (1), where $\overrightarrow{d}^i_s$ is the distance between $\overrightarrow{v}$ and the i^{th} neuron in success map, $\overrightarrow{w}^i_s$ is the weight vector of this neuron, $\overrightarrow{d}^i_s$ is the distance between $\overrightarrow{v}$ and the i^{th} neuron in failure map and $\overrightarrow{w}^i_f$ is the weight vector of this neuron. For each map, the neuron winner corresponds to the smallest distance to the vector. This distance represents therefore the distance between $\overrightarrow{v}$ and the map, see (2), where d_s is the distance between $\overrightarrow{v}$ and success map while d_f is the distance with the failure map. In the phase of decision, we compare d_s with d_f, by taking into account the threshold of vigilance s_{vig} which represents the tolerance to risks. If the threshold is higher than the difference between the distance to failure map and the distance to success map, the early warning signal becomes active, otherwise, this signal is inactive, see (3). The activation of EWS indicates that $\overrightarrow{v}$ will lead to failure if it is applied on the lower level. In the learning phase, it is possible that vector $\overrightarrow{v}$ can activate EWS at a time and inactivate it at another time because the distances with the neurons change. A decision of "nogo" corresponds to active EWS and a decision of "go" corresponds to inactive EWS. In the case where decision is "nogo", we take another vector $\overrightarrow{v}$ randomly from V, then we look for expected results by evaluation and decision phases as detailed before. In case where decision is "go" ($\overrightarrow{v}$ may lead to success), the vector will be applied on the low level controller to run a trial. After each trial there is a reward, either negative (failure) or positive (success). Only one map learns $\overrightarrow{v}$. If the reward is negative the failure map will learn, and if it is positive the success map will learn. Next, we take other vectors randomly from V and execute the same steps until the stabilization of maps. The algorithm is as follows:

1) $\cup$ ($success - map,\ failure - map$) $\in V$.
2) $\cup\ \overrightarrow{v} \in V$
   a) $Evaluation:$
   the distances to the neurons of the two maps:
$$\begin{cases} \overrightarrow{d}^i_s = -\overrightarrow{w}^i_s + \overrightarrow{v} \\ \overrightarrow{d}^i_f = -\overrightarrow{w}^i_f + \overrightarrow{v} \end{cases} \quad (1)$$
   the distances to the neurons winners of the two maps:
$$\begin{cases} d_s = min \parallel \overrightarrow{d}^i_s \parallel \\ d_f = min \parallel \overrightarrow{d}^i_f \parallel \end{cases} \quad (2)$$
   b) $Decision:$
$$EWS = \begin{cases} 0 \quad (go) \quad if (d_f - d_s) > s_{vig} \\ 1 \quad (nogo) \quad otherwise. \end{cases} \quad (3)$$
3) if ($nogo$) go to 2.
   else if (go) test $\overrightarrow{v}$, and get a reward R.
   if ($R : positive$) learn $success - map$,
   else if ($R : negative$) learn $failure - map$,
   go to 2.

D. Concept of Vigilance

Some psychological research suggest that some people are more tolerant to risk than others who are more cautious, [12][13][14]. The vigilance is related to human learning approaches and decision making [15]. In the standard psychological assessment of risk taking, people are classed as risk seeking or risk aversion [16]. In our study the vigilance is represented by a threshold s_{vig} that is used to adjust the early warning signal in the decision mechanism. This threshold describes the tolerance of risk, see Fig. 1. By definition, the threshold of vigilance is the allowed margin of the difference between the distances of state space vector $\overrightarrow{v}$ with failure map and with success map, for which the decision mechanism still responds with "go", as in (3). The threshold has a limited value according to the dimensions of the state space. In a two dimension state space $s_{vig} \in [-\sqrt{2}, +\sqrt{2}]$. Toward positive values of the threshold, the decision mechanism becomes more alert to risk (cautious). In the opposite it has a tendency to take risks (courageous), see Fig. 2, D is the diameter of the space. For instance, if $s_{vig} = 0.1$ the early warning system stays inactive for vectors closer to success map than to failure map by 0.1. In such case, if $d_s = 0.3$ and $d_f = 0.35$, then EWS becomes active, while it is inactive if $s_{vig} = 0.04$. The change in the value of s_{vig} from 0.1 to 0.04 allows the agent to be more tolerant to risks.

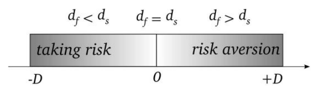


Fig. 2. Threshold of vigilance and tolerance of risk .

In this paper, we have fixed the threshold during learning, but we present the result for different values of the threshold.

III. Biological inspired neural controllers for walking

Biological inspired locomotion controllers are based on the simple circuit that is built from sensory neurons, motor neurons, and inter-neurons [9][17][18][19]. Neurophysiological studies associate the rhythmic movement with the oscillation activity of a type of neurons, called neurons oscillators [20][21]. These oscillators can produce rhythmic activity without sensory input even without central input. But the sensory information is indispensable for walking because it allows to shape the rhythmic patterns in order to interact with the environment [22]. However, sensory information are mainly used to adapt the controller in front of changes and perturbations. Neurophysiologists have proved that biological controllers like Central Pattern Generators (CPG) have an adaptation mechanism that belongs to plasticity properties [20][23]. To realize the learning approach, we are interested in having the low level controller interact with the environment, like the neural reflexive controller, proposed by [9] and tested on a real robot. This low level controller is based on the sensory motor approach. Our learning mechanism will regulate certain parameters in this controller to walk and to explore the domain of viability, that give the ability of walking adaptation to the environment.

A. Neural model for Sensory-Motor

In the neural model for sensory-motor there are direct connections between sensory neurons and motor neurons, see Fig. 3. A static model of sensory neuron has proposed by Ekberg, [18], it is described in (4), ρ_i is the activity of sensory neuron, α is a positive constant that denotes the dynamics of the neuron, θ is the amplitude and ϕ is the input on the neuron. ϕ can be an angular position, or a contact force [9]. In the other side, there is a model of motor neuron. Beer [24] has proposed a dynamic model that is described in (5), y_j is the mean membrane potential of the j^{th} motor neuron, τ is a time constant, ρ_i is the activity of the i^{th} sensory neuron, w_{ij} is the synaptic weight between the i^{th} sensory neuron and the j^{th} motor neuron, u_j is the activity of this motor neuron, θ_m is the bias.

$$\rho_i = (1 + e^{\alpha(\theta - \phi)})^{-1} \tag{4}$$

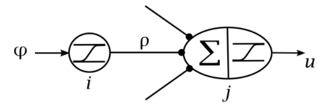


Fig. 3. A neural model of sensory motor controller.

$$\begin{cases} \tau.\frac{dy_j}{dt} = -y_j + \sum_i w_{ij}.\rho_i \\ u_j = (1 + e^{\alpha(\theta_m - y_j)})^{-1} \end{cases} \tag{5}$$

B. neural reflexive controller

The neural architecture proposed by Wörgötter to control a simulated biped [9] is based on a sensory motor approach where sensory neurons are connected to extension and flexion motor neurons. Fig. 4 shows the principles of this controller. A is a stretch receptor sensory neuron, G is a ground contact sensory neuron, FM is a flexion motor neuron, EM is an extension motor neuron. Lines with an arrow extremity indicate excitatory connections, and lines terminated by a solid circle indicate inhibitory connections. Fig. 4(a) shows the interaction between the ground contact sensory neuron of the stance leg and the flexion and extension motor neurons in this leg. Ground contact sensory neuron (G) in a leg excites the extension motor neuron (EM) in the knee and the flexion motor neuron (FM) in the hip of the same leg. Fig. 4(b) shows the interaction between ground contact sensory neuron and the flexion and extension motor neuron in the other leg. It excites the flexion motor neuron in the knee and the extension motor neuron in the hip. Fig. 4(c) shows the role of extension and flexion sensory neurons, E and F, to inhibit the corresponding motor neuron. This is the same for all joints. This behavior is referred as the articular reflex. Fig. 4(d) shows the role of the stretch receptor sensory neuron to excite the extension motor neuron in the knee of the same leg. This behavior is referred as the extension reflex.

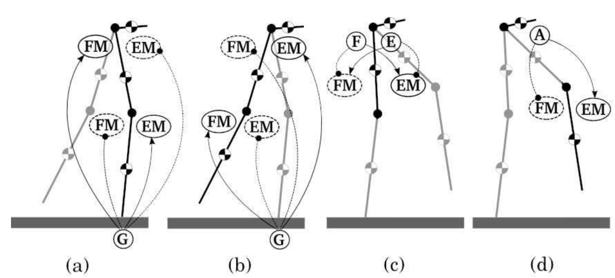


Fig. 4. Principles of the neural reflexive controller proposed by Wörgötter. (a) Interaction with the stance leg. (b) Interaction with the swing leg. (c) Articular reflex. (d) Extension reflex.

In this study, we concentrate on two parameters of this low level controller. The first α_{hip} denotes the dynamics of rhythmic movement in the hip joint (dynamics of extensor sensory neuron), while the second $\theta_{hip-max}$ represents the amplitude of this movement (amplitude in the activity of extensor sensory neuron). The biped can walk and face environment changes, such as variations in the slope of terrain , by controlling these two parameters.

C. Determination of viability domain of the neural controller

We have explored the domain of viability of the controller by varying the dynamics and the amplitude of the hip extensor sensory neuron (α_{hip} and $\theta_{hip-max}$) on a flat terrain. Inside a defined space for the two parameters, variations have been carried out with defined steps. For each couple (α_{hip},$\theta_{hip-max}$) the walking has been tested. According to definitions for success and failure we can know which couple leads to success or to failure. The biped has 10 seconds to walk, so if this time was passed and it was still staying, then it is a success. Otherwise, if it falls down before the time, it is a failure. In the simulation, we consider that the robot falls down when the gravity center of the trunk comes below the one of the two shanks. In such case the simulation will stop the trial. For all trials the robot has the same initial position in which one leg is in the stance phase and the other one is in the swing phase because we are not interested here in the initial phase of walking. Fig. 5 shows the results of this analytical studies related to walk on flat terrain. The failure trials are represented by the surrounding area, while the another area represent the success trials. α_{hip} varies in [0 : 0.5 : 20], while $\theta_{hip-max}$ varies in [90° : 1 : 150°]. Walking velocity is limited in our case between $0.33[m/s]$ and $0.66[m/s]$. In the simulation, the walking velocity corresponds to the averaged velocity measured for the trunk.

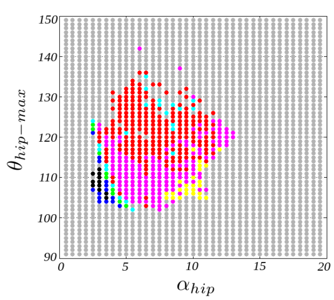


Fig. 5. Domain of viability of the low level controller in space of α_{hip} and $\theta_{hip-max}$.

IV. Learn walking and adaption approach

First, the biped will learn to walk on a flat terrain. The goal is to allow maps to explore the domain of viability in the state space. We will present the results for several values of vigilance threshold and discuss them. Second, we will present the results for a more complicated architecture devoted to learn how to walk on sloped terrain through an example that explains such adaptation approach.

A. Learn walking on flat terrain

We present how our learning approach makes success and failure maps to explore the space of parameters in order to find the domain of viability of the controller. The simulation is run for different values of vigilance threshold s_{vig}. Fig. 6 shows the control diagram in case of learning on flat terrain. There are two loops of control, a loop of low level control represented by the interaction between the biped and the controller neuronal sensory motor, the other loop concerns the high level controller where the learning mechanism controls the low level controller and receives the result for each trial (success, failure).

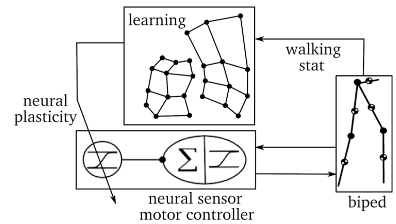


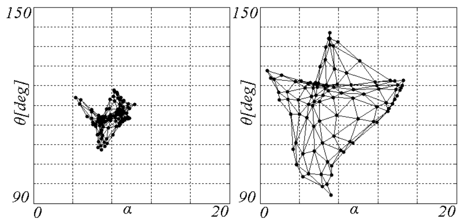
Fig. 6. Walking control diagram, composed of two control levels: neural senrory-motor controller (low level) and learning mechanism (high level).

In the learning algorithm, we initialize success map and failure map in the space of α_{hip} and $\theta_{hip-max}$. The same space has been studied previously for the domain of viability. The number of trials and the vigilance threshold are determined. For a random vector $\overrightarrow{v}(\alpha_{hip}, \theta_{hip-max})$ from this space there are two processing phases, the evaluation phase and the decision phase. If the early warning signal stays inactive for $\overrightarrow{v}$, then it may lead to success according to the past experience of this system represented by success map and failure map and also according to the risk tendency represented by vigilance threshold s_{vig}. Each vector that will lead to success has been passed to the controller sensory-motor to run a trial on the biped. According to the result of each trial one map will learn. Then, we look for another vector $\overrightarrow{v}(\alpha_{hip}, \theta_{hip-max})$ from the space, and so on. After learning, all the vectors that had led to success have been incorporated into success map and all the vectors had led to failure have been incorporated into failure map. Fig. 7 and Fig. 8 show success map and failure map after learning for 500 trials with two different threshold of vigilance. The state space is normalized between 0 and 1. Each map is composed of 100 neurons. Weights of neuron (w_1,w_2) denote a configuration of the low level controller ($w_1 = \alpha$, $w_2 = \theta_{hip-max}$). We have therefore 100 different configurations in each map that match 100 walking gaits stored in success map.

After learning with 500 trials, if $s_{vig} = 0.05$, we obtain 98% of succeeded trials, while 2% of failure. With another threshold $s_{vig} = 0$ we obtain 96% of succeeded trials, and 45% of success with $s_{vig} = -0.1$ and 28% of success with $s_{vig} = -0.2$. In the last case, as there are 72% failure, the failure map was learned better than in the other cases.

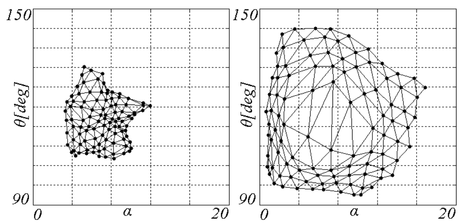
In Fig. 7 and Fig. 8 all neurons in the success map lead to success (walk), but in the second case the domain of viability presented by the zone occupied by success map is bigger than befor, which allows to have more stability and more walking gaits. So we can distinguish between two different behaviors for the system, risk taking and risk aversion.

Thanks to the two behaviors the system can gain experience in walking, but in case of risky behavior the system learns better. Fig. 9 presents the rate of success in function



(a) Success map. (b) Failure map.

Fig. 7. Success and failure maps after learning on flat terrain with vigilance threshold $s_{vig} = 0.05$.



(a) Success map. (b) Failure map.

Fig. 8. Success and failure maps after learning on flat terrain with vigilance threshold $s_{vig} = -0.2$.

of vigilance threshold, it was obtined after learning with different thresholds.

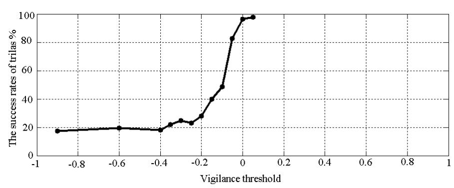
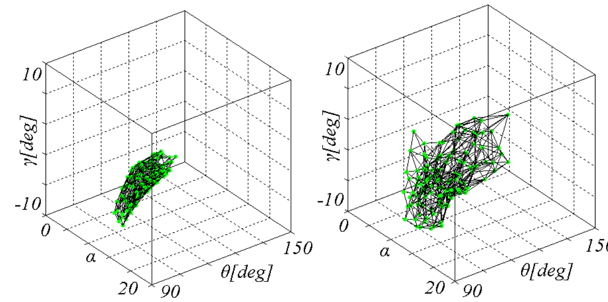


Fig. 9. Rate of succeeded trials as a function of vigilance threshold.

It could be divided into three zones. The first zone corresponds to $s_{vig} > 0.05$ where there is no decision, no trials, then no learning. The second zone corresponds to $0.05 < s_{vig} < -0.4$, the system is more risky, and for a more negative threshold the decision will be "go" for all vectors. The middle zone is the most important because it is a zone of switching between two different behaviors. In our studies we fixed the vigilance threshold during the learning phase, but changing this variable from a trial to another may be worth investigating.

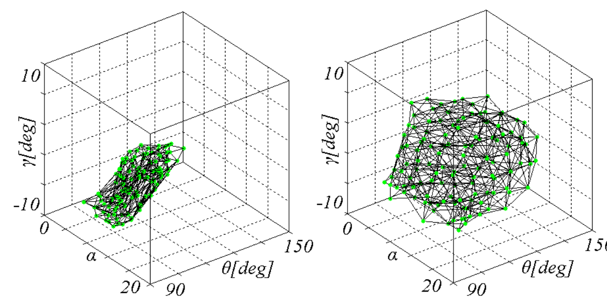
B. Learning on sloped terrain

The objective from the previous study is to represent the zone of success in the state space by success map to justify the analytical study of the domain of viability. Our objective now is to generalize the controller on sloped terrains. The modification in the maps' structures consists of adding a third dimension to describe the terrain slope γ to learn in space of α_{hip}, $\theta_{hip-max}$ and γ. In our study the slope is limited between $+10°$ and $-10°$. In the learning phase the biped learns to walk on terrains with different slopes that had been chosen randomly. After learning, the two SOM must be organized in the three dimension state space to represent success and failure experience. Fig. 10 and Fig. 11 show success and failure maps after learning for different values of vigilance threshold.



(a) Success map. (b) Failure map.

Fig. 10. Success and failure maps after learning on different terrain slopes with vigilance threshold $s_{vig} = 0.0$.



(a) Success map. (b) Failure map.

Fig. 11. Success and failure maps after learning on different terrain slopes with vigilance threshold $s_{vig} = -0.2$.

Each map is composed of 125 neurons where each neuron has three weights (w_1,w_2,w_3) that denote a configuration of the low level controller ($w_1 = \alpha_{hip}$, $w_2 = \theta_{hip-max}$) for walking on determined terrain slope ($w_3 = \gamma$). When $s_{vig} = 0$ there is a success in 86% of trials and a failure in 14%. Success and failure maps are shown in Fig. 10(a) and Fig. 10(b) respectively. For the other value of vigilance, $s_{vig} = -0.2$, there is a success only in 15% of trials and a failure in 85%, as shown in Fig. 11(a) and Fig. 11(b). The space occupied by success map in the second case is bigger than in the first case. This difference is referred to as the difference in the behavior according to vigilance threshold. As the failure rate in the second case is higher than in the first case, the failure map will learn better in the second case.

After learning, each neuron in the success map corresponds to a walking on a particular slope, including gait and speed. To walk on a terrain with a particular slope γ, a competition occurs between all neurons to find the winner without taking the (w_1, w_2) values for neurons into account. The winner is the neuron whose w_3 is the closest to γ, while other weights of the neuron winner are used to configure the parameters (α_{hip}, $\theta_{hip-max}$) of the low level controller. Changing the terrain slope during walking causes switching into another neuron in the map, which corresponds to the new slope. This switch can be direct between the neurons or indirect by use of intermediary neurons. Fig. 12 shows how the biped can walk with different slopes.

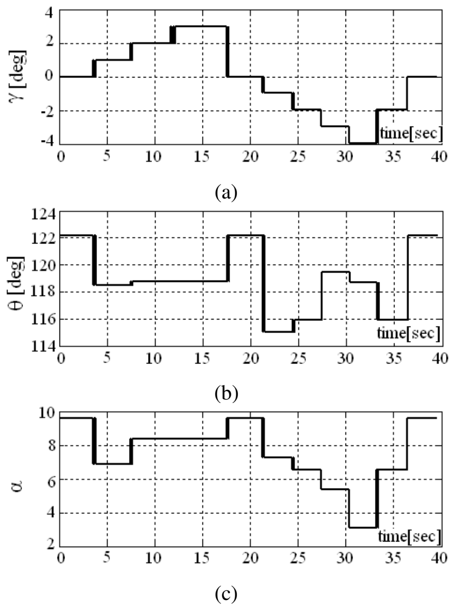


Fig. 12. Switching between the neurons of success map during walking on irregular terrain. (a) represent the terrain slope, which is an input to the learning mechanism. (b) and (c) are the amplitude and the dynamics of the extensor sensory neuron, the outputs of the learning mechanism.

For any slope γ in the domain of viability (success map) there is a corresponding couple (α_{hip}, $\theta_{hip-max}$) that can be applied to the lower level of control to perform the walking. in Fig. 12, the variation in the slope γ will cause a competition between the neurons of success map to find the more adaptive neuron with the new slope, this neuron will be able to configure the sensory-motor controller to adapt with that slope, by his weights which present the amplitude and the dynamic of the extensor sensory neuron.

V. CONCLUSION

In this paper we presented a neurobiological inspired learning algorithm. The objectives of the mechanism were to learn from mistakes and to avoid making them again. This was done by building on experience of past mistakes and successes. We showed how these two experiences could build themselves through the stages of evaluation, decision and then trials. It can be said that the negative reward has an importance as the positive. This mechanism was implemented on a planar biped and allowed the biped to learn to walk without supervision. It added the property of adaptation even to changes of terrain slope. Our future work shall address adaptation to further changes in the environment, as well as changes in the physical parameters of the biped, as an important factor to allow our mechanism to apply to a real robot. An analytical study that will investigate how much benefit the system is able to get from previous experience will also be considered.

REFERENCES

[1] T. Singer, B. Seymour, J. O'Doherty, H. Kaube, R.J. Dolan and C.D. Frith, Empathy for pain involves the affective but not sensory components of pain, *Science*, vol. 303. no. 5661, 2004, pp. 1157-1162.

[2] J.W. Brown and T.S. Braver, A computational model of risk, conflict, and individual difference effects in the anterior cingulate cortex, *Brain Research*, vol. 1202, 2008, pp.99-108.

[3] RB. Mars, MG. Coles, MJ.Grol, CB.Holroyd, S.Nieuwenhuis, W.Hulstijn and I.Toni, Neural dynamics of error processing in medial frontal cortex, *NeuroImage*, vol. 28, no. 4, 2005, pp. 1007-1013.

[4] H. Gemba, K. Sasaki and V.B. Brooks, Error potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neuroscience letters*, 1986, vol. 70, no. 2, pp. 223-227.

[5] W.J. Gehring, M.G.H. Coles, D.E. Meyer and E. Donchin, The error-related negativity: An event-related potential accompanying errors, *Psychophysiology*, 1990, vol. 27, S34.

[6] J. Hohnsbein, M. Falkenstein and J. Hoorman, Error processing in visual and auditory choice reaction tasks, *Journal of Psychophysiology*, 1989, vol. 3, 32.

[7] J.W. Brown and T.S. Braver, Learned Predictions of Error Likelihood in the Anterior Cingulate Cortex *Science*, 2005, vol. 307, no. 5712, pp. 1118-1121.

[8] L. Van Leijenhorst, P.M. Westenberg and E.A. Crone, A developmental study of risky decisions on the Cake Gambling Task: Age and gender analyses of probability estimation and reward evaluation, *Developmental Neuropsychology*, 2008, vol. 33, no. 2, pp. 179-196.

[9] T. Geng, B. Porr and F. Wörgötter, Fast Biped Walking with a Sensor-driven Neuronal Controller and Real-time Online Learning, *The International Journal of Robotics Research*, 2006, vol. 25, no. 3, pp. 243-259.

[10] K. Matsumoto, W. Suzuki and K. Tanaka, Neuronal correlates of goal-based motor selection in the prefrontal cortex, *Science*, 2003, vol. 301, no. 5630, pp. 229-232.

[11] T. Kohonen, Self-Organizing Maps, *Springer Series in Information Sciences*, vol. 30, Springer, Berlin, Heidelberg, New York, 1995, 1997, 2001. Third Extended Edition, 501 pages. ISBN 3-540-67921-9, ISSN 0720-678X

[12] J.-L. van Gelder, R.E. de Vries and J. van der Plight, Evaluating a dual-process model of risk: affect and cognition as determinants of risky choice, *Journal of Behavioral Decision Making*, 2008, vo. 22, no. 1, pp. 45-61.

[13] B. Pawlowski, R. Atwal and R.I.M. Dunbar, Sex differences in everyday risk-taking behavior in humans, *Evolutionary Psychology*, 2008, vol.6, no.1, pp. 29-42.

[14] P. Horvarth and M. Zuckerman, Sensation seeking, risk appraisal, and risky behavior. *Personality and individual differences*, 1993, vol. 14, no. 1, pp. 41-52

[15] H. Ahn, and R. Picard, Affective-Cognitive Learning and Decision Making: A Motivational Reward Framework For Affective Agent, *The 1st International Conference on Affective Computing and Intelligent Interaction*. October 22-24, 2005, Beijing, China.

[16] D.J. Kruger, X.T. Wang, and A. Wilke, Towards the development of an evolutionarily valid domain-specific risk-taking scale, *Evolutionary Psychology*, 2007, vol. 5, no. 3, pp. 555-568.

[17] G. Taga, Y. Yamaguchi and H. Shimizu, Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment, *Biological Cybernetics*, 1991, vol. 65, no.3, pp. 147-159.

[18] T. Wadden and O. Ekeberg, A neuro-mechanical model of legged locomotion: single leg control, *Biological Cybernetics*, 1998, vol. 79, no.2, pp. 161-173.

[19] H. Cruse, Ch. Bartling, M. Dreifert, J. Schmitz, D.E. Brunn, J. Dean and T. Kindermann, Walking: a complex behavior controlled by simple systems, *Adaptive Behavior*, 1995, vol. 3, no. 4, pp. 385-418.

[20] D.A. McCrea and I.A. Rybak, Organization of mammalian locomotor rhythm and pattern generation, *Brain Research Reviews*, 2008, vol. 57, no. 1, pp. 134-146.

[21] P.F. Rowat and A.I. Selverston, Learning algorithms for oscillatory networks with gap junctions and membrane currents, *Network*, 1991, vol. 2, no. 1, pp. 17-41.

[22] E. Marder and RL. Calabrese, Principles of rhythmic motor pattern generation, *Physiological Reviews*, 1996, vol. 76, no. 3, pp. 687717.

[23] A. Ishiguro, A. Fujii and P.E. Hotz, Neuromodulated Control of Bipedal Locomotion Using a Polymorphic CPG Circuit, *Adaptive Behavior*, 2003, vol. 11, no. 1, pp. 7-17.

[24] R.D. Beer, H.J. Chiel, R.D. Quinn, K.S. Espenschied and P. Larsson, A distributed neural network architecture for hexapod robot locomotion, *Neural Computation*, 1992, vol. 4, no. 3, pp. 356-365.